中文引用格式:牛莉霞,林彦宏. 人机信任对协同决策质量的影响:人机共享心智模式的中介作用[J]. 中国安全科学学报,2026,36(2):244-252.

英文引用格式:NIU Lixia, LIN Yanhong. Impact of human-AI trust on collaborative decision-making quality: mediating role of human-AI shared mental model[J]. China Safety Science Journal, 2026, 36(2): 244-252.

人机信任对协同决策质量的影响:
人机共享心智模式的中介作用*

牛莉霞 教授, 林彦宏**
(辽宁工程技术大学 工商管理学院,辽宁 葫芦岛 125105)

中图分类号:X915.2 **文献标志码**:A **DOI**: 10.16265/j.cnki.issn1003-3033.2026.02.1791
基金项目:国家自然科学基金资助(52174184)。

【摘 要】 为提升复杂工业场景下人机协同决策质量,缓解人类对人工智能(AI)的信任不足并弥合人机协作中的认知差异,基于心智理论,构建人机信任、人机共享心智模式(Human-AI SMM)与人机协同决策质量之间的关系模型,并引入任务复杂性作为调节变量。首先,根据各变量间的理论关系提出假设,并结合人机信任量表、Human-AI SMM 量表、人机协同决策质量量表以及任务复杂性量表设计问卷;然后,向全国范围内 AI 使用企业的一线员工发放问卷,并收集有效样本 493 份;最后,采用 SPSS 26.0、AMOS 24.0 及 Process 4.0 对收集的有效样本进行数据分析与假设检验。结果表明:人机信任显著正向影响人机协同决策质量;Human-AI SMM 在人机信任与人机协同决策质量之间起中介作用;任务复杂性正向调节人机信任与 Human-AI SMM 之间的关系。

【关键词】 人机信任; 人机协同决策质量; 人机共享心智模式(Human-AI SMM); 心智理论; 中介作用; 任务复杂性

Impact of human-AI trust on collaborative decision-making quality:
mediating role of human-AI shared mental model

NIU Lixia, LIN Yanhong
(School of Business Administration, Liaoning Technical University, Huludao Liaoning 125105, China)

Abstract: In order to enhance human-artificial intelligence(AI) collaborative decision-making quality in complex industrial environments, mitigate the deficiency of human trust in AI, and bridge cognitive gaps in human-AI collaboration, based on the theory of mind, the study constructed a model of the relationship between human-AI trust, Human-AI SMM and human-AI collaborative decision-making quality, and task complexity was introduced as a moderating variable. First, hypotheses were proposed based on the theoretical relationships among the variables, and a questionnaire was designed by integrating the human-AI trust scale, Human-AI SMM scale, the human-AI collaborative decision-making quality scale, and the task complexity scale. Then, the questionnaires were distributed to frontline employees of AI-using companies nationwide, and 493 valid samples were collected. Finally, SPSS 26.0, AMOS 24.0 and

* 文章编号:1003-3033(2026)02-0244-09; 收稿日期:2025-08-20; 修稿日期:2025-11-29
** 通信作者:林彦宏(2001—),女,辽宁铁岭人,硕士研究生,研究方向为安全人因工程和组织行为管理。E-mail: 2892465729@qq.com。

Process 4.0 were used for data analysis and hypothesis testing on the collected valid samples. The results of the study show that human-AI trust significantly and positively affects human-AI collaborative decision-making quality. Human-AI SMM mediates the relationship between human-AI trust and human-AI collaborative decision-making quality. Task complexity positively moderates the relationship between human-AI trust and Human-AI SMM.

Keywords: human-AI trust; human-AI collaborative decision-making quality; human-AI shared mental model (Human-AI SMM); theory of mind; mediating role; task complexity

0 引　言

在现代复杂的工业场景中,人工智能(Artificial Intelligence,AI)技术的深度参与为人机协同决策提供了强大的技术支持。人类的灵活性与有限的认知能力,能够与机器强大的数据处理能力形成高度互补[1],这种协同效应使人类与 AI 协同完成任务的决策效率与质量往往优于单独完成任务[2]。因此,人机协同决策模式不仅能够提升决策速度与质量,还有利于降低人为失误的风险,显著提高复杂工业环境中的生产效率与安全性。

目前,国内外学者针对信任在人机协同决策中的关键作用,已展开重要探索,例如:KEDING 等[3]从信任的视角评估了基于 AI 的咨询系统如何影响人类的决策行为,发现信任能够显著提高决策过程的感知质量。孔祥维等[4]从模型、数据和信任等方面提出人机协同可信决策框架,并对 AI 使能系统可信决策研究进行了展望。ASAN 等[5]分析了 AI 系统在医疗协作中的应用效果,指出人机协作的有效性不仅取决于 AI 系统背后的底层数学过程的准确性,更受包括信任在内的人为因素影响。何贵兵等[6]的研究指出,人类决策者对智能体的信任是人机协同决策得以开展的前提。DE PACE 等[7]探讨了人机协作中人与机器人的交流问题,指出员工的信任影响后续人机协作意图,从而对人机协作产生至关重要的影响。但是,当前研究较少从认知科学和心智理论的视角深入分析人机信任的形成和发展过程,对于信任背后的人机心智过程的认识仍显不足。

鉴于此,笔者拟从心智理论的视角出发,系统性地分析人机信任对人机协同决策质量的影响机制,特别关注人机共享心智模式(Human-AI Shared Mental Model,Human-AI SMM)的重要性及任务复杂性的调节作用,以期为提升人机协同决策质量提供新的研究视角,同时,为企业在复杂多变的环境中优化人机协作模式提供理论指导。

1 心智理论基础与研究假设

1.1 人机信任对人机协同决策质量的影响假设

人机信任是指个体在不确定或脆弱的情况下,对依赖 AI 系统帮助其实现目标的态度[8],以及对 AI 不会利用己方弱点的信心和心理预期[9]。人机协同决策质量是指在人类与 AI 共同参与决策过程中,所达到的决策结果的准确性、可靠性和效果[10]。根据心智理论,AI 通过模拟人类的心智过程,可以更准确地预测人类的行为,并提供适应性响应,甚至成为团队协作中的有效伙伴[11]。在人机团队协作中,具备心智理论的 AI 能够识别人类成员的角色与责任,理解其工作负荷和情绪状态,并提供适时地帮助和支持,从而增强人类对 AI 的信任。这种信任关系在协同决策过程中起着关键作用。一方面,这种信任增强人类对 AI 反馈的认可度,使他们更愿意采纳 AI 提供的信息和建议。研究表明:人们对于 AI 越信任,就会越愿意采纳 AI 系统[12],从而利用其反馈提升人机协同决策的精准度和可靠性。另一方面,这种信任让人类将 AI 视为具有“成长型心态”的协作伙伴,并增强与其持续互动和磨合的意愿。通过不断的交互与磨合,AI 能够更准确地理解人类需求,人类也能更熟悉 AI 的工作逻辑,从而提升人机协同决策质量。因此,提出假设:

H_1:人机信任对人机协同决策质量有显著正向影响。

1.2 Human-AI SMM 的中介作用假设

SMM 指的是团队成员之间对于任务目标、操作程序、角色分配以及策略方法的共同理解和认识[13]。在以 AI 为协作伙伴的人机团队中,Human-AI SMM 是指人类与 AI 对于任务目标、操作程序、角色分配以及策略方法的共同理解和认识,反映了人机在协作过程中的认知一致性。研究表明:在传统团队中,团队成员之间的信任有利于 SMM 的形成[14]。类似的,在人机协作环境下,人对 AI 的信任

也可能促进 SMM 的建立。然而,与人际信任不同,这里的信任是单向的,即指人对 AI 的信任,而非“人机互信”。根据心智理论,AI 可以更有效地以人类可理解的方式进行沟通,就像人类能够做的那样[11],使员工能够更好的理解和推断 AI 的行为和意图,从而建立对 AI 的信任。例如:员工会认为,具备心智理论的 AI 不仅能完成任务,还具备“共情”与“理解”其工作意图的能力,从而更愿意与 AI 共享任务相关的深层次信息。这种信任驱动的信息共享与认知一致性,为 Human-AI SMM 的建立提供了基础。

CANNON-BOWERS 等[15]研究发现,当团队成员拥有相似的态度或信念时,他们会对任务或环境产生一致的解释,从而作出有效决策。根据心智理论,具备心智能力的 AI 通过理解团队成员的心理状态和能力,能够更准确地预测和适应人类行为[11]。当人机之间建立起 SMM 时,AI 不仅能够提升对人类行为的适应性,还能够增强人类对 AI 行为的理解性与预测性。这种心智模式使得人和 AI 能够更准确地理解彼此的意图和任务需求,从而减少信息不对称与误解,提升协同决策的科学性与一致性。因此,提出假设:

H_2:Human-AI SMM 在人机信任与人机协同决策质量之间起中介作用。

1.3 任务复杂性的调节作用假设

任务复杂性是指工作任务的复杂程度,它受到多样性、相互依赖性、不确定性等多种因素的影响,同时也受任务执行者的认知能力和经验等主观性因素的影响[16]。通过模拟人类的心智理论,AI 可以更准确地预测人类的行为,提供更合适的响应,甚至参与到团队合作中,成为协作伙伴[11]。当面向某些复杂任务时,AI 已经具备了足够的能力承担特定的重要任务,并在与人类组成人机混合协作群体时,通过不同的协同模式增强群体性能[17]。在高任务复杂性的情境中,人类认知负荷的增加使其更可能依赖 AI 系统的分析与支持,与 AI 的交互也会更加频繁和深入。通过不断的互动和反馈,从而加速双方对彼此能力与局限的理解,进一步加强 Human-AI SMM 的构建。相比之下,在低任务复杂性的情景中,任务要求较低。这些任务往往处于他们的舒适区,所需的技能也较为基础,他们无需依赖额外的技术支持或复杂的知识系统即可完成[18]。此时,人类对 AI 的依赖性较弱,与 AI 的互动也较少,Human-AI SMM 的增强程度较为有限。因此,提出假设:

H_3:任务复杂性在人机信任与 Human-AI SMM 之间起正向调节作用。

基于以上分析,构建人机信任、任务复杂性、Human-AI SMM 与人机协同决策质量之间的关系模型,如图 1 所示。

图 1 研究假设模型

Fig. 1 Research hypothesis model

2 一线员工人机信任问卷调查

2.1 人机信任问卷调查对象与调查过程

以全国范围内 AI 使用企业的一线员工为研究对象,样本来源涵盖制造业、高科技行业、医疗保健行业等多个领域。为确保数据的真实性和有效性,正式调研前向参与者详细说明研究目的,明确告知该调研仅用于科学研究,不涉及任何商业用途或个人利益。同时,所有问卷均采取匿名填写方式,并承诺对收集到的所有数据严格保密,不会泄露参与者的任何个人信息,从而确保参与者能够真实表达其观点与态度。问卷通过问卷星平台在线发放,共向 AI 使用企业的一线员工发放问卷 579 份。为保证数据质量,在问卷回收后对数据进行严格筛选,剔除答案高度一致、逻辑矛盾、漏答或多选的无效问卷。最终得到有效问卷 493 份,有效问卷回收率为 85.15%。在所有被调查的企业员工样本中,性别以男性为主,占比为 63.5%;年龄主要集中在 26~35 岁,占比为 46.7%;教育程度本科学历占比较高,为 52.1%;所属行业是制造业、医疗保健行业和能源行业占比较大;工作年限以 1~5 年为主,占比为 52.7%。

2.2 人机信任及 Human-AI SMM 量表选择

为保证变量测量可靠性和有效性,选用国内外使用较频繁的相关成熟量表,其中,有关英文量表均按照严格的翻译-回译程序进行本土化完善,并结合研究情境对部分题项进行适当修订。除人口统计学变量外,其他测量量表均采用李克特 5 点计分法,1—5 表示“完全不符合”到“完全符合”,分值越高表示符合程度越高。具体测量如下:

1) 人机信任。采用 GULATI 等[19]改编的量表,结合文中情境进行修订,共包含 12 个题项,如“我认为使用 AI 可能会产生负面后果”等。通过数据分析,该量表的 Cronbach′s Alpha 系数为 0.849。

2) Human-AI SMM。参考王黎萤等[20]编制的协作式共享心智模型量表和项凯标[21]的研究成果。结合研究情境,对不适用的题项进行删减和修订,最终确定 9 个题项,如“我认为我与 AI 所具有的不同专门知识都是完成任务所需要的”等。通过数据分析,该量表的 Cronbach′s Alpha 系数为 0.926。

3) 人机协同决策质量。基于 SHAMIM 等[22]开发的决策质量量表,对不适用于本次测量情境的题项进行了删减。结合文中情境进行修订,共包含 6 个题项,如“我认为我与 AI 一起作出了好的工作决策”等。通过分析,该量表的 Cronbach′s Alpha 系数为 0.883。

4) 任务复杂性。参考 KUHN 等[23]开发的量表,并结合研究情境进行调整和修订,最终量表包含 6 个题项,如“我认为完成我与 AI 协同进行的决策任务需要付出很多精力”等。通过分析,该量表的 Cronbach′s Alpha 系数为 0.886。

5) 控制变量。结合以往研究,将员工的性别、年龄、学历、工作年限等人口统计学变量作为控制变量,以排除其对研究结果可能产生的干扰。此外,由于不同行业对 AI 的应用程度不同,可能对员工的行为和认知产生一定影响,因此将所属行业作为控制变量纳入模型分析。

2.3　量表数据处理及分析方法

采用 SPSS 26.0 和 AMOS 24.0 软件统计分析样本数据。首先,采用 Harman 单因素分析进行检验评估样本是否存在共同方法偏差,然后采用验证性因子分析检验量表的区分效度;其次,对样本的基本特征进行描述性统计分析,并通过相关分析检验各变量之间的初步关系;最后,通过分层回归对研究假设进行验证,分析人机信任对人机协同决策质量的直接效应、Human-AI SMM 的中介效应及任务复杂性的调节效应,并通过简单斜率分析进一步验证任务复杂性在模型中的调节作用。

3　人机信任数据结果与分析

3.1　验证性因子分析与区分效度检验

运用 AMOS 24.0 软件对人机信任、任务复杂性、Human-AI SMM 和人机协同决策质量进行验证性因子分析,来检验变量的区分效度,结果见表 1。

表 1　验证性因子分析结果

Table 1　Results of confirmatory factor analysis

模型	χ^2	df	χ^2/df	CFI	TLI	RMSEA	SRMR
4 因子模型(A、B、C、D)	878.694	489	1.797	0.948	0.944	0.040	0.044
3 因子模型(A、B+C、D)	2427.923	492	4.935	0.741	0.723	0.089	0.103
2 因子模型(A+B+C、D)	3647.496	494	7.384	0.579	0.550	0.114	0.137
单因子模型(A+B+C+D)	4878.705	495	9.856	0.415	0.376	0.134	0.158
加入共同方法因子	706.726	457	1.546	0.967	0.961	0.033	0.040

注:A 表示人机信任,B 表示 Human-AI SMM,C 表示任务复杂性,D 表示人机协同决策质量,+表示合并。χ^2 表示卡方统计量;自由度(Degree of Freedom,df);比较适配指数(Comparative Fit Index,CFI);测试理论指数(Test of Logical Interpretation,TLI);近似误差均方根(Root Mean Square Error of Approximation,RMSEA);标准化根均方残差(Standardized Root Mean Square Residua,SRMR)。

由表 1 可知:4 因子模型是各种模型中拟合最佳的 (χ^2/df = 1.797,CFI = 0.948,TLI = 0.944,RMSEA = 0.040,SRMR = 0.044),这说明各变量区分效度良好,可以进行后续变量间关系的检验。

为检验共同方法偏差,首先,利用 SPSS 26.0 软件对数据进行 Harman 单因素分析,结果显示,第 1 个析出因子的因子载荷量为 23.261%,未超过 40%这一建议值,并有 6 个因子特征值大于 1,初步可以判定共同方法偏差不严重。其次,利用 AMOS 24.0 软件在 4 因子模型的基础之上加入 1 个共同方法因子,再次检验共同方法偏差问题。在 4 因子模型中加入共同方法因子之后,模型拟合指标并未得到显著改善(χ^2/df = 1.546,CFI = 0.967,TLI = 0.961,RMSEA = 0.033,SRMR = 0.040),ΔCFI 和 ΔTLI 未超过 0.1,ΔRMSEA 和 ΔSRMR 未超过 0.05,进一步表明不存在严重的共同方法偏差。

3.2　各变量描述性统计与相关分析

主要变量的平均值、标准差及变量间相关系数见表 2,由表 2 可知:人机信任与 Human-AI SMM 显著正相关($r = 0.286, p < 0.01$),与人机协同决

策质量显著正相关($r = 0.189, p < 0.01$);Human-AI SMM 与人机协同决策质量显著正相关($r = 0.218, p < 0.01$)。这些相关性与理论预期一致,因此,研究假设得到初步支持。

表 2 描述性统计与相关分析结果($N=493$)

Table 2 Results of descriptive statistics and correlation analysis ($N=493$)

变量	均值	标准差	1	2	3	4	5	6	7	8	9
性别	1.37	0.482	1	—	—	—	—	—	—	—	—
年龄	1.86	0.774	−0.006	1	—	—	—	—	—	—	—
学历	2.16	0.674	−0.078	−0.143**	1	—	—	—	—	—	—
所属行业	3.65	1.809	0.065	−0.047	−0.011	1	—	—	—	—	—
.工作年限	1.69	0.920	−0.002	0.310**	0.015	−0.021	1	—	—	—	—
人机信任	2.965	0.87807	−0.065	0.038	−0.052	0.030	0.166**	1	—	—	—
Human-AI SMM	3.014	1.113	0.114*	−0.055	−0.032	0.068	0.096*	0.286**	1	—	—
任务复杂性	2.694	1.051	0.001	0.100*	−0.024	0.054	0.044	−0.172**	−0.166**	1	—
人机协同决策质量	3.041	1.122	0.024	−0.026	−0.008	−0.084	0.030	0.189**	0.218**	−0.212**	1

注:* 表示 $p<0.05$,** 表示 $p<0.01$,*** 表示 $p<0.001$,下同。

3.3 Human-AI SMM 的中介作用验证

1)主效应检验。采用层次回归分析方法分析人机信任、Human-AI SMM 和人机协同决策质量之间的关系,结果见表 3。由模型 4 可知:人机信任显著正向影响人机协同决策质量(回归系数 $\beta = 0.249, p < 0.001$)。由此,假设 H_1 得到验证。

2)中介效应检验。在表 3 中,由模型 2 可知:人机信任显著正向影响 Human-AI SMM ($\beta = 0.357, p < 0.001$)。由模型 5 可知:Human-AI SMM 显著正向影响人机协同决策质量($\beta = 0.223, p < 0.001$)。模型 6 在模型 4 基础上加入中介变量 Human-AI SMM,效应值显著($\beta = 0.183, p < 0.001$),人机信任对人机协同决策质量效应值显著($\beta = 0.184, p < 0.01$)。

使用软件 SPSS26.0 中 Process4.0 宏程序,设定置信区间为 95%,通过 5 000 次偏倚矫正的 bootstrap 抽样,在控制性别、年龄、学历、所属行业和工作年限的情况下,对 Human-AI SMM 在人机信任和人机协同决策质量之间的中介效应进行检验,结果见表 4。由表 4 可知:人机信任对人机协同决策质量影响的直接效应为 0.184, bootstrap95% 置信区间为[0.068, 0.300]及 Human-AI SMM 的中介效应为 0.065, bootstrap95% 置信区间为[0.028, 0.106], bootstrap95% 置信区间的上、下限均不包含 0,表明人机信任不仅能够直接影响人机协同决策质量,而且能够通过 Human-AI SMM 的中介作用影响人机协同决策质量。由此,假设 H_2 得到验证。

表 3 Human-AI SMM 的中介效应回归分析

Table 3 Regression analysis of mediating effect of Human-AI SMM

类别		Human-AI SMM		人机协同决策质量			
		模型 1	模型 2	模型 3	模型 4	模型 5	模型 6
控制变量	性别	0.247*	0.294**	0.066	0.099	0.011	0.045
	年龄	−0.139*	−0.129	−0.066	−0.059	−0.035	−0.036
	学历	−0.064	−0.035	−0.022	−0.002	−0.008	0.004
	所属行业	0.036	0.030	−0.054	−0.058*	−0.062*	−0.064*
	工作年限	0.155**	0.095	0.052	0.010	0.017	−0.007
自变量	人机信任	—	0.357***	—	0.249***	—	0.184**
中介变量	Human-AI SMM	—	—	—	—	0.223***	0.183***
R^2		0.035	0.112	0.011	0.047	0.048	0.076
ΔR^2		0.035	0.077	0.011	0.037	0.047	0.029
F		3.529**	10.171***	1.037	4.008**	4.984***	5.725***
ΔF		3.529**	41.898***	1.037	18.671***	24.470***	15.318***

3)调节效应检验。根据调节效应检验办法检验任务复杂性在人机信任与 Human-AI SMM 之间的调节作用。为避免多重共线性,对人机信任与任务复杂性进行中心化处理,结果见表 5。表 5 中,在控制相关变

量后,由模型 9 可知:人机信任与任务复杂性的交乘项与 Human-AI SMM 之间为显著正向相关关系(β = 0.111,p < 0.05),说明任务复杂性正向调节人机信任与 Human-AI SMM 之间的关系,因此,假设 3 得到验证。

表 4　总效应、直接效应及中介效应分析

Table 4　Analysis of total effect, direct effect and mediation effect

效应	效应值	标准误差	95%置信区间
中介效应	0.065	0.020	[0.028, 0.106]
直接效应	0.184	0.059	[0.068, 0.300]
总效应	0.249	0.058	[0.136, 0.362]

表 5　任务复杂性调节效应检验

Table 5　Moderating effect test of task complexity

类别		Human-AI SMM		
		模型 7	模型 8	模型 9
控制变量	性别	0.247*	0.290**	0.278**
	年龄	−0.139*	−0.113	−0.111
	学历	−0.064	−0.039	−0.041
	所属行业	0.036	0.035	0.036
	工作年限	0.155**	0.102	0.096
主效应	人机信任	—	0.329***	0.340***
	任务复杂性	—	−0.128**	−0.115*
调节效应	人机信任×任务复杂性	—	—	0.111*
R^2		0.035	0.125	0.135
ΔR^2		0.035	0.090	0.009
F		3.529**	9.938***	9.437***

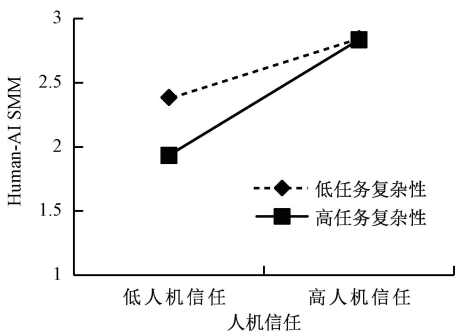

图 2　不同任务复杂性水平下人机信任与 Human-AI SMM 之间的关系

Fig. 2　Relationship between human-AI trust and Human-AI SMM at different levels of task complexity

为了能更加清晰地表明任务复杂性在人机信任与 Human-AI SMM之间的调节作用,通过简单斜率分析绘制调节效应图,如图 2 所示,由图 2 可知:当任务复杂性较高时,人机信任对 Human-AI SMM 的正向作用较强,即任务复杂性正向调节人机信任与 Human-AI SMM 的正向关系。由此,假设 H_3 得到支持。

为进一步检验调节效应,分析当任务复杂性处于低、高不同水平时,人机信任和 Human-AI SMM 的关系,检验结果见表 6,由表 6 可知:任务复杂性处于低水平时,人机信任对 Human-AI SMM 的影响效应估计值为 0.223,95% 置信区间为[0.081, 0.365],不包含 0;任务复杂性处于高水平时,人机信任对 Human-AI SMM 的影响效应估计值为 0.456,95%置信区间为[0.303,0.610],不包含 0,说明任务复杂性显著调节人机信任对 Human-AI SMM 的影响。假设 3 得到支持。

表 6　调节效应(Bootstrap 法) 检验结果

Table 6　Test results for moderating effects(Bootstrap)

任务复杂性	效应值	标准误差	95%置信区间	
			低限	高限
低任务复杂性	0.223	0.072	0.081	0.365
高任务复杂性	0.456	0.078	0.303	0.610

4 人机信任对协同决策质量的影响分析

4.1 人机信任对人机协同决策质量的影响

人机信任显著正向影响人机协同决策质量,进一步验证了信任是高质量协同决策的前提条件。在复杂工业环境中,安全问题往往是决策过程中至关重要的因素。AI 凭借强大的数据处理能力和客观性,能够减少因人类情绪波动或认知偏差引发的错误决策,这对于降低安全隐患尤为重要。当员工信任 AI 时,他们更愿意将部分或全部任务交由 AI 处理,从而避免因自身判断失误或偏见导致的风险。尤其在安全任务中,AI 能够提供更精确、实时的决策支持,显著提升决策质量。例如:在智能生产线的安全管理中,AI 能够预测设备故障、监测环境变量并采取预防措施,降低事故发生的概率。

此外,人机信任还能够提升决策效率。在高风险环境下,AI 通过实时监控和数据分析,能够快速发现潜在隐患并提供反馈。当操作员信任 AI 时,他们会更依赖系统的实时反馈,而非反复验证或质疑,从而减少决策过程中的拖延,提升整体决策效率与可靠性。因此,企业应高度重视员工对 AI 的信任建设,通过增强 AI 的透明性与可靠性,缓解员工的不信任感,从而充分释放人机协同的潜力,提升人机协同决策质量并降低安全风险。

4.2 Human-AI SMM 的中介作用

基于心智理论验证了 Human-AI SMM 在人机信任与人机协同决策质量之间的中介作用,进一步揭示了人机信任对协同决策质量影响的内在机制。在高风险工业场景中,人对 AI 的信任能够促进信息的高效交互与共享,进而形成 SMM。为了更好地达到 Human-AI SMM,企业应该提升人类对 AI 的信任,这不仅要提高 AI 的透明度和可解释性,还需确保其稳定性、可靠性与安全性,以减少使用者的不确定性与顾虑;同时,发展 AI 的心智能力也很重要,通过增强其情境理解与意图推理能力,使其更好地适应人类的认知模式,从而强化人机协作的共享心智。如此,形成的 SMM 有助于人类和 AI 更准确地预测对方的行为,减少信息不对称与理解偏差,从而提升协同决策质量。

具体而言,Human-AI SMM 通过提升人机认知一致性优化协同效能。在工业场景(如生产线、危险作业)中,AI 通过实时传感器数据分析,快速识别故障或风险,而人类基于对 AI 行为模式的理解进行精准响应,这种 SMM 能显著提升突发事件中的决策效率,降低因认知偏差导致的失误。此外,Human-AI SMM 还能够增强团队的适应性,特别是在面对动态变化的任务环境时。由于事件发展的偶然性、高度不确定性以及人们认知的有限性,使得决策者难以把握事故发生的条件或发展态势[24]。人机基于共享的认知框架,能够在任务执行过程中灵活调整策略,确保决策过程的连贯性与高效性。因此,企业在推动人机协同时,应注重构建 SMM,通过信任建设与信息共享优化人机协同过程,提升协同决策质量。

4.3 任务复杂性的调节作用

任务复杂性在人机信任与 Human-AI SMM 之间起正向调节作用。在高任务复杂性情境下,人类决策者因认知能力与经验的局限性,更倾向于依赖 AI 的支持。这种依赖增强了人机之间的互动与信息共享,促进了 Human-AI SMM 的形成与稳定。高复杂性任务中,AI 系统能够通过快速处理海量数据、实时反馈和预测功能,有效弥补人类的认知不足,尤其在安全管理领域表现尤为突出。智能系统能够有效预测潜在的安全风险,为安全执行提供行为指导[25]。例如:在工业生产过程中,AI 系统能够实时监测设备状态、环境变化及安全标准,并及时发出预警,辅助人类决策者进行风险管理与应急反应。这种人机协作在一定程度上减少了因人类认知局限性导致的错误判断,从而确保任务的高效执行与安全性。

相比之下,在低任务复杂性情境下,人类能够凭借自身知识与经验轻松完成任务,因而对 AI 的依赖较低,与 AI 的交互也相对较少,Human-AI SMM 的形成受到一定限制。这一发现表明:人机信任对 Human-AI SMM 的影响因任务复杂性的强弱而表现出差异性,企业在设计人机协同框架时,应根据任务复杂性的不同,动态调整人机协作模式,确保协同效能的最大化。

5 结 论

1) 人机信任与人机协同决策质量显著正相关,人机信任与 Human-AI SMM 显著正相关,Human-AI SMM 与人机协同决策质量显著正相关。

2) 人机信任能够通过 Human-AI SMM 影响人机协同决策质量;任务复杂性在人机信任与 Human-AI SMM 之间起正向调节作用。

3) 尽管问卷调查法在测量个体主观感知方面具有优势,但可能存在自陈数据的主观偏差。未来

研究可结合试验法或神经科学方法(如眼动追踪、脑电分析),进一步验证变量间的因果关系。此外,目前大多数研究聚焦于人类对 AI 的信任,而较少关注 AI 对人类的信任和适应能力。未来研究可以进一步拓展这一方向,为高风险、高复杂性任务环境中的决策优化提供更强的支持。

参 考 文 献

[1] BOYACI T, CANYAKMAZ C, DE VÉRICOURT F. Human and machine: the impact of machine input on decision making under cognitive limitations[J]. Management Science, 2024, 70(2): 1 258-1 275.

[2] FÜGENER A, GRAHL J, GUPTA A, et al. Cognitive challenges in human-artificial intelligence collaboration: investigating the path toward productive delegation[J]. Information Systems Research, 2022, 32(2): 678-696.

[3] KEDING C, MEISSNER P. Managerial overreliance on AI-augmented decision-making processes: how the use of AI-based advisory systems shapes choice behavior in R&D investment decisions[J]. Technological Forecasting and Social Change, 2021, 171: DOI:10. 1016/j. techfore. 2021. 120970.

[4] 孔祥维,王子明,王明征,等. 人工智能使能系统的可信决策:进展与挑战[J]. 管理工程学报,2022,36(6):1-14.
KONG Xiangwei, WANG Ziming, WANG Mingzheng, et al. Trustworthy decision-making in artificial intelligence-enabled systems: progress and challenges[J]. Journal of Industrial Engineering and Engineering Management, 2022, 36(6): 1-14.

[5] ASAN O, BAYRAK A E, CHOUDHURY A. Artificial intelligence and human trust in healthcare: focus on clinicians[J]. Journal of Medical Internet Research, 2020, 22(6): DOI:10. 2196/15154.

[6] 何贵兵,陈诚,何泽桐,等. 智能组织中的人机协同决策:基于人机内部兼容性的研究探索[J]. 心理科学进展, 2022, 30(12): 2 619-2 627.
HE Guibing, CHEN Cheng, HE Zetong, et al. Human-agent collaborative decision-making in intelligent organizations: a perspective of human-agent inner compatibility[J]. Advances in Psychological Science, 2022, 30(12): 2 619-2 627.

[7] DE PACE F, MANURI F, SANNA A, et al. A systematic review of augmented reality interfaces for collaborative industrial robots[J]. Computers & Industrial Engineering, 2020, 149: DOI:10. 1016/j. cie. 2020. 106806.

[8] LEE J D, SEE K A. Trust in automation: designing for appropriate reliance[J]. Human Factors, 2004, 46(1): 50-80.

[9] 黄心语,李晔. 人机信任校准的双途径:信任抑制与信任提升[J]. 心理科学进展, 2024, 32(3): 527-542.
HUANG Xinyu, LI Ye. Trust dampening and trust promoting: a dual-pathway of trust calibration in human-robot interaction[J]. Advances in Psychological Science, 2024, 32(3): 527-542.

[10] LATINOVIC Z, CHATTERJEE S C. Achieving the promise of AI and ML in delivering economic and relational customer value in B2B[J]. Journal of Business Research, 2022, 144: 966-974.

[11] BENDELL R, WILLIAMS J, FIORE S M, et al. Individual and team profiling to support theory of mind in artificial social intelligence[J]. Scientific Reports, 2024, 14(1): DOI:10. 1038/s41598-024-63122-8.

[12] 吴俊,张迪,刘涛,等. 人类对人工智能信任的接受度及脑认知机制研究:实证研究与神经科学实验的元分析[J]. 管理工程学报, 2024, 38(1): 60-73.
WU Jun, ZHANG Di, LIU Tao, et al. A study on the acceptance of human trust in artificial intelligence and brain cognitive mechanism: a meta-analysis of empirical studies and neuroscience experiments[J]. Journal of Industrial Engineering and Engineering Management, 2024, 38(1): 60-73.

[13] 喻国明,陈雪娇. 人机协同下合作行为的发生机制与演变路径[J]. 山东师范大学学报:社会科学版,2024,69(1): 146-156,165.
YU Guoming, CHEN Xuejiao. Occurrence mechanism and evolution path of cooperative behavior in human-autonomy teaming[J]. Journal of Shandong Normal University: Social Sciences, 2024, 69(1): 146-156, 165.

[14] 武欣,吴志明. 团队共享心智模型的影响因素与效果[J]. 心理学报, 2005, 37(4): 542-549.
WU Xin, WU Zhiming. Antecedents and consequences of shared mental model in work teams[J]. Acta Psychologica Sinica, 2005, 37(4): 542-549.

[15] CANNON-BOWERS J A, SALAS E. Reflections on shared cognition[J]. Journal of Organizational Behavior: The

International Journal of Industrial, Occupational and Organizational Psychology and Behavior, 2001, 22(2): 195-202.

[16] CAMPBELL D J. Task complexity: a review and analysis[J]. Academy of Management Review, 1988, 13(1): 40-52.

[17] SHIRADO H, CHRISTAKIS N A. Locally noisy autonomous agents improve global human coordination in network experiments[J]. Nature, 2017, 545(7654): 370-374.

[18] ZHU Yi, WANG Taotao, WANG Chang, et al. Complexity-driven trust dynamics in human-robot interactions: insights from AI-enhanced collaborative engagements[J]. Applied Sciences, 2023, 13(24): DOI:10.3390/app132412989.

[19] GULATI S, SOUSA S, LAMAS D. Design, development and evaluation of a human-computer trust scale[J]. Behaviour &information Technology, 2019, 38(10): 1 004-1 015.

[20] 王黎萤,陈劲. 研发团队创造力的影响机制研究:以团队共享心智模型为中介[J]. 科学学研究, 2010, 28(3): 420-428.
WANG Liying, CHEN Jin. A study on influencing mechanism of R&D team creativity: the mediating role of team shared mental model[J]. Studies in Science of Science, 2010, 28(3): 420-428.

[21] 项凯标. 动态环境下团队过程、共享心智模型和组织绩效关系研究[D]. 北京:北京交通大学, 2014.
XIANG Kaibiao. The research on relationships of team process, shared mental model and organizational performance in dynamic environment[D]. Beijing: Beijing Jiaotong University, 2014.

[22] SHAMIM S, ZENG J, SHARIQ S M, et al. Role of big data management in enhancing big data decision-making capability and quality among Chinese firms: a dynamic capabilities view[J]. Information & Management, 2019, 56: DOI:10.1016/j.im.2018.12.003.

[23] KUHN T, POOLE M S. Do conflict management styles affect group decision making? evidence from a longitudinal field study[J]. Human Communication Research, 2000, 26(4): 558-590.

[24] 孙美兰,邹树梁,徐守龙,等. 基于动态贝叶斯网络的乏燃料后处理核应急情景分析[J]. 中国安全科学学报, 2024, 34(10): 214-220.
SUN Meilan, ZOU Shuliang, XU Shoulong, et al. Nuclear emergency scenario analysis for spent fuel reprocessing based on dynamic Bayesian network[J]. China Safety Science Journal, 2024, 34(10): 214-220.

[25] 牛莉霞,李肖萌,王洪妹. 5G 时代负系统学视域下智能系统安全分析体系构建[J]. 中国安全科学学报, 2024, 34(8): 1-10.
NIU Lixia, LI Xiaomeng, WANG Hongmei. Construction of intelligent system security analysis system under view of negative systematics in 5G era[J]. China Safety Science Journal, 2024, 34(8): 1-10.

作者简介: 牛莉霞 (1983—),女,山西吕梁人,博士,教授,主要从事组织行为管理与人因工程方面的研究。E-mail:nlx8941@126.com。